CLAIMS

We claim:

1.   A computer-implemented method for identifying sequences of Web
pages that are frequently visited in order non-consecutively during user browsing
sessions, the Web pages of a sequence visited in order during a user browsing session
when each Web page of the sequence is visited before a next Web page of the sequence is
visited, the Web pages of a sequence visited non-consecutively during a user browsing
session when at least one intervening Web page that is not part of the sequence is visited
between the visits to the Web pages of the sequence, the method comprising:

receiving a Web server log generated by a Web server serving a Website
having Web pages, the Web server log reflecting usage of the Website by users;

identifying multiple user browsing sessions from the received Web server
log, each identified user browsing session indicating a series of Web pages from the
Website that were consecutively visited in order by a user; and

identifying sequences of the Web pages of the Website that were frequently
visited in order non-consecutively in the identified user browsing sessions by

determining a minimum threshold number of user browsing sessions;

determining multiple Web pages that are each visited during more of
the identified user browsing sessions than the minimum threshold;

creating a tree data structure having a root node and having a child
node of the root node for each of the determined Web pages, each of the determined Web
pages represented by one of the children nodes, the children nodes forming a current
lowest level of the tree data structure;

repeatedly expanding the tree data structure by adding a new lowest
level of nodes that are children nodes to the nodes of a previous lowest level of the tree
data structure, the added children nodes such that each of the determined Web pages has
a node that represents that determined Web page that is added as a child node to each of

26 the nodes of the previous level, each of the added children nodes having an associated

27 sequence of Web pages consisting of the determined Web pages that are represented by

28 the nodes in a path from the root node to that node;

29   determining the nodes of the tree data structure whose associated

30 sequence of Web pages is visited in order non-consecutively during more of the identified

31 user browsing sessions than the minimum threshold; and

32   removing the nodes of the tree data structure that are not among the

33 determined nodes,

34 such that after creation of the tree data structure is completed, the sequences of Web

35 pages that are associated with the nodes remaining in the tree data structure are the

36 identified sequences of the Web pages that were frequently visited in order non-

37 consecutively during the identified user browsing sessions.

1   2.   The method of claim 1 including producing a report that includes the

2 identified frequently visited sequences of Web pages.

1   3.   The method of claim 1 wherein the acts of determining the nodes

2 and removing the nodes are each performed after each act of adding a new lowest level of

3 nodes, and wherein the act of determining the nodes includes:

4   initializing counters associated with each of the nodes other than the root

5 node;

6   for each of the identified user browsing sessions, incrementing the counter

7 associated with each node whose associated sequence of Web pages is visited in order

8 non-consecutively during the identified user browsing session; and

9   after the incrementing of the counters for each of the identified user

10 browsing sessions, selecting the nodes of the tree data structure whose associated

11 counters have values greater than the minimum threshold to be the determined nodes.

1   4.   The method of claim 3 including creating for each of the determined

2 Web pages a linked list data structure that is associated with that determined Web page,

3   and wherein the incrementing of the counters associated with the nodes for a user

4   browsing session includes:

5   for each of the linked list data structures, initializing the linked list to include

6   an entry for the child node of the root node that represents the determined Web page with

7   which the linked list is associated; and

8   selecting in order each Web page in the series indicated by the user

9   browsing session and, if that Web page is one of the determined Web pages,

10   incrementing the counter associated with each of the nodes having

11   entries in the linked list data structure associated with the one determined Web page;

12   selecting all children nodes of each node having an entry in the linked

13   list data structure associated with the one determined Web page;

14   removing each of the entries in the linked list data structure associated

15   with the one determined Web page; and

16   for each of the selected children nodes, adding an entry for the

17   selected child node to the linked list data structure associated with the determined Web

18   page represented by the selected child node.

1   5.   The method of claim 3 wherein the incrementing of the counters for

2   an identified user browsing session includes:

3   associating a token with the root node; and

4   selecting in order each Web page in the series indicated by the user

5   browsing session and, if that Web page is one of the determined Web pages,

6   for each node with an associated token, determining any children

7   nodes of that node that represent the one determined Web page; and

8   for each determined child node, if the node does not have an

9   associated token,

10   incrementing the counter associated with the determined child

11   node; and

12   associating a token with the determined child node.

1   6.   The method of claim 3 including, after the adding of each new
2 lowest level of nodes but before the determining of the nodes, removing the added nodes
3 of the new lowest level whose associated sequence of Web pages includes a subsequence
4 of Web pages that is not associated with any of the nodes that were part of the tree data
5 structure before the adding.

1   7.   The method of claim 3 wherein the repeated expansion of the tree
2 data structure is ended when all of the nodes of an added lower level are removed.

1   8.   The method of claim 1 wherein the repeated expansion of the tree
2 data structure is ended when a maximum number of levels of nodes is reached.

1   9.   The method of claim 1 including, after the determining of the
2 multiple Web pages, determining sequences of two of the determined Web pages that are
3 visited in order non-consecutively during more of the identified user browsing sessions
4 than the minimum threshold, wherein during the adding of each new lowest level of
5 nodes an added child node is removed if any consecutive two of the Web pages in the
6 sequence of Web pages associated with that child node are not one of the determined
7 sequences of two determined Web pages, and wherein the repeated expanding of the tree
8 data structure is completed before the determining of the nodes.

1   10.   The method of claim 9 wherein the determining of the sequences of
2 two of the determined Web pages includes creating an array data structure having entries
3 representing each possible sequence of two of the determined Web pages.

1   11.   The method of claim 1 wherein the determining of the minimum
2 threshold number of user browsing sessions is based on an indication received from a
3 user.

1   12.   A computer-implemented method for identifying frequent sequences

2 of interaction events, the method comprising:

3   from at least one interaction log that contains data reflecting interactions

4 with at least one executing application, identifying multiple interaction sessions each

5 indicating an ordered series of related interaction events; and

6   identifying sequences of interaction events that frequently occur in the

7 identified interaction sessions by

8   determining multiple interaction events that are each present in at least

9 a first threshold number of the identified interaction sessions;

10   for each of the determined interaction events, identifying a sequence

11 consisting of that interaction event as being a sequence of interaction events of length 1

12 that frequently occurs, the length 1 being a current longest length of identified sequences;

13 and

14   repeatedly identifying sequences of interaction events of increasing

15 lengths that frequently occur by,

16   for each of the identified sequences of a current longest length,

17 generating candidate sequences of interaction events by adding one of the determined

18 interaction events at the end of the identified sequence;

19   identifying the newly generated candidate sequences that are

20 present in at least a second threshold number of the identified interaction sessions as

21 being sequences of interaction events that frequently occur, each of the identified newly

22 generated sequences having a current longest length that is one greater than the previous

23 longest length; and

24   if none of the newly generated candidate sequences are

25 identified as sequences of interaction events that frequently occur, ending the repeated

26 identification of the sequences.

1   13.   The method of claim 12 wherein the identifying of the sequences of

2 interaction events of length 1 includes creating a data structure with a root element and

3   multiple children elements of the root element, each child element representing one of the

4   determined interaction events and being associated with the identified sequence of length

5   1 that consists of that one determined interaction event, the children elements forming a

6   current lowest level of the data structure, and wherein the generating of candidate

7   sequences of a length one greater than the current longest length includes expanding the

8   data structure by adding a new lowest level of elements to the data structure such that the

9   added elements are children elements to the elements of a previous lowest level of the

10   data structure, each of the added children elements representing one of the determined

11   interaction events and having an associated sequence of interaction events that is one of

12 the generated candidate sequences.

1   14.   The method of claim 13 wherein each element other than the root

2   element has an associated sequential path of elements between the root element and that

3   element, a first element in each sequential path being a child element of the root element,

4   each element in each sequential path other than the first element being a child element of

5   the previous element in the sequential path, and wherein the sequence of interaction

6   events that is associated with each node other than the root node consists of a sequence of

7   the interaction events represented by the elements in the path associated with that node

8   followed by the interaction event represented by that node.

1   15.   The method of claim 13 wherein the identifying of the newly

2   generated candidate sequences as sequences of interaction events that frequently occur

3   includes determining the newly added children elements whose associated sequence of

4   interaction events is present in at least the second threshold number of the identified

5   interaction sessions.

1   16.   The method of claim 15 wherein the determining of the newly added

2   children elements whose associated sequence of interaction events is present in at least

3   the second threshold number of the identified interaction sessions includes, for each of

4   the identified interaction sessions:

5 initializing for each of the determined interaction events a linked list data

6 structure that is associated with that determined interaction event such that the linked list

7 data structure includes an entry for the child element of the root element that represents

8 the determined interaction event with which the linked list is associated; and

9 selecting in order each interaction event in the identified interaction session

10 and, if that interaction event is one of the determined interaction events,

11 indicating for each of the elements having entries in the linked list data

12 structure associated with the one determined interaction event that the sequence of

13 interaction events associated with the element is present in the identified interaction

14 session;

15 selecting all children elements of each element having an entry in the

16 linked list data structure associated with the one determined interaction event;

17 removing each of the entries in the linked list data structure associated

18 with the one determined interaction event; and

19 for each of the selected children elements, adding an entry for the

20 selected child element to the linked list data structure associated with the determined

21 interaction event represented by the selected child element.

1 17. The method of claim 15 wherein the determining of the newly added

2 children elements whose associated sequence of interaction events is present in at least

3 the second threshold number of the identified interaction sessions includes, for each of

4 the identified interaction sessions:

5 associating a token with the root element; and

6 selecting in order each interaction event in the identified interaction session

7 and, if that interaction event is one of the determined interaction events,

8 for each element with an associated token, determining any children

9 elements of that element that represent the one determined interaction event; and

10 for each determined child element, if the element does not have an

11 associated token,

12   indicating that the sequence of interaction events associated with

13   the determined child element is present in the identified interaction session; and

14   associating a token with the determined child element.

1   18.   The method of claim 15 wherein the identifying of the newly

2   generated candidate sequences as sequences of interaction events that frequently occur

3   includes, after the determining of the newly added children elements whose associated

4   sequence of interaction events is present in at least a second threshold number of the

5   identified interaction sessions, removing the newly added children elements that are not

6   among the determined elements.

1   19.   The method of claim 12 wherein the second threshold number is the

2   first threshold number.

1   20.   The method of claim 12 wherein the generating of the candidate

2   sequences is such that a candidate sequence of interaction events is not generated if the

3   candidate sequence includes a subsequence of interaction events that is not one of the

4   identified sequences of interaction events that frequently occurs.

1   21.   The method of claim 12 including:

2   determining a maximum sequence length for the identified sequences of

3   interaction events that frequently occur; and

4   ending the repeated identification of the sequences after the identifying of

5   the newly generated candidate sequences that have a current longest length that is the

6   specified maximum sequence length.

1   22.   The method of claim 12 including providing indications of the

2   identified sequences of interaction events that frequently occur to at least one user.

1   23.   A computer-readable medium containing instructions that when
2 executed cause a computing device to identify frequent sequences of interaction events
3 by:

4   identifying multiple interaction sessions each indicating an ordered series of
5 related interaction events; and

6   identifying sequences of interaction events that frequently occur in the
7 identified interaction sessions by

8   determining multiple interaction events that are each present in at least
9 a first threshold number of the identified interaction sessions;

10   for each of the determined interaction events, identifying a sequence
11 consisting of that interaction event as being a sequence of interaction events of length 1
12 that frequently occurs, the length 1 being a current longest length of identified sequences;
13 and

14   repeatedly identifying sequences of interaction events of increasing
15 lengths that frequently occur by,

16   for each of the identified sequences of a current longest length,
17 generating candidate sequences by adding one of the determined interaction events at the
18 end of the identified sequence;

19   identifying the newly generated candidate sequences that are
20 present in at least a second threshold number of the identified interaction sessions as
21 being sequences of interaction events that frequently occur, each of the identified newly
22 generated sequences having a current longest length that is one greater than the previous
23 longest length; and

24   if none of the newly generated candidate sequences are
25 identified as sequences of interaction events that frequently occur, ending the repeated
26 identification of the sequences.

1   24.   The computer-readable medium of claim 23 wherein the identifying
2 of the sequences of interaction events of length 1 includes creating a data structure with a

3 root element and multiple children elements of the root element, each child element
4 representing one of the determined interaction events and being associated with the
5 identified sequence of length 1 that consists of that one determined interaction event, the
6 children elements forming a current lowest level of the data structure, and wherein the
7 generating of candidate sequences of a length one greater than the current longest length
8 includes expanding the data structure by adding a new lowest level of elements to the
9 data structure such that the added elements are children elements to the elements of a
10 previous lowest level of the data structure, each of the added children elements
11 representing one of the determined interaction events and having an associated sequence
12 of interaction events that is one of the generated candidate sequences.

1 25.   The computer-readable medium of claim 24 wherein the identifying
2 of the newly generated candidate sequences as sequences of interaction events that
3 frequently occur includes determining the newly added children elements whose
4 associated sequence of interaction events is present in at least the second threshold
5 number of the identified interaction sessions by, for each of the identified interaction
6 sessions:

7   initializing for each of the determined interaction events a linked list data
8 structure that is associated with that determined interaction event such that the linked list
9 data structure includes an entry for the child element of the root element that represents
10 the determined interaction event with which the linked list is associated; and

11   selecting in order each interaction event in the identified interaction session
12 and, if that interaction event is one of the determined interaction events,

13   indicating for each of the elements having entries in the linked list data
14 structure associated with the one determined interaction event that the sequence of
15 interaction events associated with the element is present in the identified interaction
16 session;

17   selecting all children elements of each element having an entry in the
18 linked list data structure associated with the one determined interaction event;

19   removing each of the entries in the linked list data structure associated

20 with the one determined interaction event; and

21   for each of the selected children elements, adding an entry for the

22 selected child element to the linked list data structure associated with the determined

23 interaction event represented by the selected child element.

1   26.   The computer-readable medium of claim 24 wherein the identifying

2 of the newly generated candidate sequences as sequences of interaction events that

3 frequently occur includes determining the newly added children elements whose

4 associated sequence of interaction events is present in at least the second threshold

5 number of the identified interaction sessions by, for each of the identified interaction

6 sessions:

7   associating a token with the root element; and

8   selecting in order each interaction event in the identified interaction session

9 and, if that interaction event is one of the determined interaction events,

10   for each element with an associated token, determining any children

11 elements of that element that represent the one determined interaction event; and

12   for each determined child element, if the element does not have an

13 associated token,

14   indicating that the sequence of interaction events associated with

15 the determined child element is present in the identified interaction session; and

16   associating a token with the determined child element.

1   27.   The computer-readable medium of claim 23 wherein the instructions

2 further cause the computing device to indicate the identified sequences of interaction

3 events that frequently occur to at least one user.

1   28.   A computing device for identifying frequent sequences of interaction

2 events, comprising:

3   an interaction event supplier component capable of identifying multiple

4 interaction sessions each indicating an ordered series of related interaction events from at

5 least one interaction log; and

6   a frequent sequence analyzer component capable of identifying sequences of

7 interaction events that frequently occur in the identified interaction sessions by

8   determining multiple interaction events that are each present in at least

9 a first threshold number of the identified interaction sessions;

10   for each of the determined interaction events, identifying a sequence

11 consisting of that interaction event as being a sequence of interaction events of length 1

12 that frequently occurs, the length 1 being a current longest length of identified sequences;

13 and

14   repeatedly identifying sequences of interaction events of increasing

15 lengths that frequently occur by,

16   for each of the identified sequences of a current longest length,

17 generating candidate sequences by adding one of the determined interaction events at the

18 end of the identified sequence;

19   identifying the newly generated candidate sequences that are

20 present in at least a second threshold number of the identified interaction sessions as

21 being sequences of interaction events that frequently occur, each of the identified newly

22 generated sequences having a current longest length that is one greater than the previous

23 longest length; and

24   if none of the newly generated candidate sequences are

25 identified as sequences of interaction events that frequently occur, ending the repeated

26 identification of the sequences.

1   29.   The computing device of claim 28 including a frequent sequence

2 indication generator component capable of providing to at least one user indications of

3 the identified sequences of interaction events that frequently occur.

1   30.   The computing device of claim 28 wherein the identifying of the

2 sequences of interaction events of length 1 includes creating a data structure with a root

3 element and multiple children elements of the root element, each child element

4 representing one of the determined interaction events and being associated with the

5 identified sequence of length 1 that consists of that one determined interaction event, the

6 children elements forming a current lowest level of the data structure, and wherein the

7 generating of candidate sequences of a length one greater than the current longest length

8 includes expanding the data structure by adding a new lowest level of elements to the

9 data structure such that the added elements are children elements to the elements of a

10 previous lowest level of the data structure, each of the added children elements

11 representing one of the determined interaction events and having an associated sequence

12 of interaction events that is one of the generated candidate sequences.

1   31.   The computing device of claim 30 wherein the identifying of the

2 newly generated candidate sequences as sequences of interaction events that frequently

3 occur includes determining the newly added children elements whose associated

4 sequence of interaction events is present in at least the second threshold number of the

5 identified interaction sessions by, for each of the identified interaction sessions:

6   initializing for each of the determined interaction events a linked list data

7 structure that is associated with that determined interaction event such that the linked list

8 data structure includes an entry for the child element of the root element that represents

9 the determined interaction event with which the linked list is associated; and

10   selecting in order each interaction event in the identified interaction session

11 and, if that interaction event is one of the determined interaction events,

12   indicating for each of the elements having entries in the linked list data

13 structure associated with the one determined interaction event that the sequence of

14 interaction events associated with the element is present in the identified interaction

15 session;

16   selecting all children elements of each element having an entry in the

17 linked list data structure associated with the one determined interaction event;

18   removing each of the entries in the linked list data structure associated

19 with the one determined interaction event; and

20   for each of the selected children elements, adding an entry for the

21 selected child element to the linked list data structure associated with the determined

22 interaction event represented by the selected child element.

1   32.   The computing device of claim 30 wherein the identifying of the

2 newly generated candidate sequences as sequences of interaction events that frequently

3 occur includes determining the newly added children elements whose associated

4 sequence of interaction events is present in at least the second threshold number of the

5 identified interaction sessions by, for each of the identified interaction sessions:

6   associating a token with the root element; and

7   selecting in order each interaction event in the identified interaction session

8 and, if that interaction event is one of the determined interaction events,

9   for each element with an associated token, determining any children

10 elements of that element that represent the one determined interaction event; and

11   for each determined child element, if the element does not have an

12 associated token,

13   indicating that the sequence of interaction events associated with

14 the determined child element is present in the identified interaction session; and

15   associating a token with the determined child element.

1   33.   A computer-implemented method for identifying frequent sequences

2   of interaction events, the method comprising:

3   from at least one interaction log that contains data reflecting interactions

4   with at least one executing application, identifying multiple interaction sessions each

5   indicating an ordered series of related interaction events; and

6   identifying sequences of interaction events that frequently occur in the

7   identified interaction sessions by

8   determining sequences of interaction events each having a specified

9   number of multiple interaction events such that the determined sequences are each

10   present in at least a first threshold number of the identified interaction sessions;

11   generating candidate sequences of interaction events having varying

12   lengths, each generated candidate sequence such that each subsequence of the candidate

13   sequence that is of a length that is the specified number is one of the determined

14   sequences, by adding at least one interaction event at the end of each of the determined

15   sequences and by repeatedly generating additional sequences by adding at least one

16   interaction event at the end of previously generated sequences; and

17   identifying the generated candidate sequences that are present in at

18   least a second threshold number of the identified interaction sessions as being sequences

19   of interaction events that frequently occur.

1   34.   The method of claim 33 wherein the generating of the candidate

2   sequences of interaction events includes creating a multi-level data structure having a root

3   element and a plurality of other elements, a first of the levels of the data structure having

4   elements that are children elements of the root element, each of the other levels having

5   elements that are children elements of elements of a previous level, and each of the other

6   elements representing one of the interaction events and being associated with one of the

7   generated candidate sequences.

35. The method of claim 34 wherein each element other than the root element has an associated sequential path of elements between the root element and that element, a first element in each sequential path being a child element of the root element, each element in each sequential path other than the first element being a child element of the previous element in the sequential path, and wherein the sequence of interaction events that is associated with each node other than the root node consists of a sequence of the interaction events represented by the elements in the path associated with that node followed by the interaction event represented by that node.

36. The method of claim 34 wherein the identifying of the generated candidate sequences as being sequences of interaction events that frequently occur includes determining the elements of the data structure whose associated sequence of interaction events is present in at least the second threshold number of the identified interaction sessions.

37. The method of claim 36 wherein the determining of the elements of the data structure whose associated sequence of interaction events is present in at least the second threshold number of the identified interaction sessions includes, for each of the identified interaction sessions:

for each interaction event represented by a child element of the root element, initializing a linked list data structure that is associated with that interaction event such that the linked list data structure includes an entry for the child element of the root element that represents the interaction event with which the linked list is associated; and

selecting in order each interaction event in the identified interaction session and, if that interaction event is one of the interaction events having an associated linked list data structure,

indicating for each of the elements having entries in the associated linked list data structure that the sequence of interaction events associated with that element is present in the identified interaction session;

15   selecting all children elements of each element having an entry in the

16   associated linked list data structure;

17   removing each of the entries in the associated linked list data structure;

18   and

19   for each of the selected children elements that represent an interaction

20   event that has an associated linked list data structure, adding an entry for the selected

21   child element to that linked list data structure.

1   38.   The method of claim 36 wherein the determining of the elements of

2   the data structure whose associated sequence of interaction events is present in at least the

3   second threshold number of the identified interaction sessions includes, for each of the

4   identified interaction sessions:

5   associating a token with the root element; and

6   selecting in order each interaction event in the identified interaction session

7   and for each selected interaction event,

8   for each element with an associated token, determining any children

9   elements of that element that represent the selected interaction event; and

10   for each determined child element, if the element does not have an

11   associated token,

12   indicating that the sequence of interaction events associated with

13   the determined child element is present in the identified interaction session; and

14   associating a token with the determined child element.

1   39.   The method of claim 36 wherein the identifying of the generated

2   candidate sequences as being sequences of interaction events that frequently occur

3   includes, after the determining of the elements of the data structure whose associated

4   sequence of interaction events is present in at least the second threshold number of the

5   identified interaction sessions, removing the elements of the data structure that are not

6   among the determined elements.

1   40.   The method of claim 33 wherein the second threshold number is the

2   first threshold number.

1   41.   The method of claim 33 including:

2   determining a maximum sequence length for the identified sequences of

3   interaction events that frequently occur; and

4   ending the repeated generation of the additional sequences after the

5   generating of candidate sequences whose length is the specified maximum sequence

6   length.

1   42.   The method of claim 33 including providing indications of the

2   identified sequences of interaction events that frequently occur to at least one user.

1   43.   A computer-readable medium containing instructions that when

2   executed cause a computing device to identify frequent sequences of interaction events

3   by:

4   from at least one interaction log that contains data reflecting interactions

5   with at least one executing application, identifying multiple interaction sessions each

6   indicating an ordered series of related interaction events; and

7   identifying sequences of interaction events that frequently occur in the

8   identified interaction sessions by

9   determining sequences of interaction events each having a specified

10   number of multiple interaction events such that the determined sequences are each

11   present in at least a first threshold number of the identified interaction sessions;

12   generating candidate sequences of interaction events of varying

13   lengths, each generated candidate sequence such that each subsequence of the candidate

14   sequence that is of a length that is the specified number is one of the determined

15   sequences, by adding at least one interaction event at the end of each of the determined

16 sequences and by repeatedly generating additional sequences by adding at least one

17 interaction event at the end of previously generated sequences; and

18 identifying the generated candidate sequences that are present in at

19 least a second threshold number of the identified interaction sessions as being sequences

20 of interaction events that frequently occur.

1 44.   The computer-readable medium of claim 43 wherein the generating

2 of the candidate sequences of interaction events includes creating a multi-level data

3 structure having a root element and a plurality of other elements, a first of the levels of

4 the data structure having elements that are children elements of the root element, each of

5 the other levels having elements that are children elements of elements of a previous

6 level, and each of the other elements representing one of the interaction events and being

7 associated with one of the generated candidate sequences.

1 45.   The computer-readable medium of claim 43 wherein the instructions

2 further cause the computing device to indicate the identified sequences of interaction

3 events that frequently occur to at least one user.

1 46.   A computing device for identifying frequent sequences of interaction

2 events, comprising:

3 an interaction event supplier component capable of identifying multiple

4 interaction sessions each indicating an ordered series of related interaction events from at

5 least one interaction log; and

6 a frequent sequence analyzer component capable of identifying sequences of

7 interaction events that frequently occur in the identified interaction sessions by

8 determining sequences of interaction events each having a specified

9 number of multiple interaction events such that the determined sequences are each

10 present in at least a first threshold number of the identified interaction sessions;

11 generating candidate sequences of interaction events of varying

12 lengths, each generated candidate sequence such that each subsequence of the candidate

13 sequence that is of a length that is the specified number is one of the determined

14 sequences, by adding at least one interaction event at the end of each of the determined

15 sequences and by repeatedly generating additional sequences by adding at least one

16 interaction event at the end of previously generated sequences; and

17 identifying the generated candidate sequences that are present in at

18 least a second threshold number of the identified interaction sessions as being sequences

19 of interaction events that frequently occur.

1 47. The computing device of claim 46 including a frequent sequence

2 indication generator component capable of providing to at least one user indications of

3 the identified sequences of interaction events that frequently occur.

1 48. The computing device of claim 46 wherein the generating of the

2 candidate sequences of interaction events includes creating a multi-level data structure

3 having a root element and a plurality of other elements, a first of the levels of the data

4 structure having elements that are children elements of the root element, each of the other

5 levels having elements that are children elements of elements of a previous level, and

6 each of the other elements representing one of the interaction events and being associated

7 with one of the generated candidate sequences.

1 49. A computer-implemented method for identifying ordered sequences

2 of usage events that frequently occur during usage sessions, the method comprising:

3 for each of multiple usage sessions, receiving an indication of an ordered

4 sequence of usage events that occurred during the usage session;

5 for each of multiple sequences of usage events, determining that the

6 sequence of usage events is present in the indicated usage sessions more times than a

7 threshold, a sequence of usage events determined to be present in a usage session when

8 each usage event of the sequence has a distinct matching event in the ordered sequence

9 for the usage session such that the matching event for each usage event in the sequence

10   appears in the ordered sequence before the matching event for a next usage event in the

11   sequence; and

12   providing indications of the determined sequences of usage events.

1   50.   The method of claim 49 including receiving at least one log from a

2   customer reflecting usage by multiple users of a service provided by that customer,

3   wherein the indications of the ordered sequences of related usage events for the multiple

4   usage sessions are received by identifying the multiple usage sessions from the received

5   log, and wherein the indications of the determined sequences of usage events are

6   provided to a representative of the customer.

1   51.   The method of claim 50 wherein the provided indications are part of

2   a generated report.

1   52.   The method of claim 49 including receiving a remote instruction

2   from a user, and wherein the determining that the sequences of usage events are present

3   more than the threshold is in response to the received instruction.

1   53.   The method of claim 49 including determining the threshold based

2   on an indication received remotely from a user.

1   54.   The method of claim 49 including receiving an instruction from a

2   remote user, and presenting the determined sequences of usage events to the remote user

3   in response.

1   55.   The method of claim 54 wherein the presenting includes providing a

2   Web page to the remote user that includes interactive indications of the determined

3   sequences of usage events.

1   56.   The method of claim 49 wherein the multiple sequences of usage

2   events are each present in usage sessions in a non-consecutive order in at least some of

3   the more times, a sequence of usage events determined to be present in a usage session in

4   a non-consecutive order when two of the matching events for usage events of the

5   sequence are separated in the ordered sequence of the usage session by at least one

6   intervening usage event that is not part of the sequence of usage events.

1   57.   The method of claim 56 wherein the determining that each of

2   multiple sequences of usage events is present more than the threshold includes:

3   determining multiple usage events that are each present in at least a

4   minimum number of the indicated usage sessions;

5   generating candidate sequences of usage events each consisting of at least

6   one of the determined usage events; and

7   repeatedly

8   selecting each of the current candidate sequences that is present in

9   more of the indicated usage sessions than the minimum number as one of the determined

10   sequences of usage events; and

11   generating additional candidate sequences of usage events by adding at

12   least one of the determined usage events to the end of at least some of the selected

13   candidate sequences.

1   58.   The method of claim 56 wherein the determining that each of

2   multiple sequences of usage events is present more than a threshold includes:

3   determining multiple usage events that are each present in at least a

4   minimum number of the indicated usage sessions;

5   determining combinations of two of the determined usage events such that

6   each combination is present in at least the minimum number of the identified interaction

7   sessions;

8   generating a plurality of candidate sequences of usage events such that each

9   subsequence of two usage events in each candidate sequence is one of the determined

10   combinations; and

11   selecting each of the generated candidate sequences that is present in more

12   of the indicated usage sessions than the minimum number as being one of the determined

13   sequences of usage events.

1   59.   The method of claim 58 wherein each of the determined

2   combinations of two determined usage events is a sequence of the two usage events.

1   60.   The method of claim 49 wherein the determining that each of

2   multiple sequences of usage events is present more than the threshold includes:

3   determining multiple usage events that are each present in at least a

4   minimum number of the indicated usage sessions;

5   when a number of the determined usage events is sufficiently small,

6   determining the multiple sequences by

7   generating candidate sequences of usage events of varying lengths;

8   and

9   validating all of the generated candidate sequences with a single

10   reading of each of the indicated usage sessions; and

11   when the number of the determined usage events is not sufficiently small,

12   determining the multiple sequences by repeatedly

13   generating groups of candidate sequences of usage events based on

14   shorter sequences of usage events that were previously validated; and

15   validating the candidate sequences of the most recently generated

16   group by reading each of the indicated usage sessions.

1   61.   The method of claim 49 wherein a sequence of usage events is

2   determined to be present in a usage session only if the matching event in that usage

3   session for each usage event of the sequence is separated from the matching event in that

4   usage session for the next usage event of the sequence by less than a determined number

5   of other usage events in the usage session.

1   62.   The method of claim 49 wherein each usage event in each usage

2   session has an associated time of occurrence, and wherein a sequence of usage events is

3   determined to be present in a usage session only if the occurrence time of the matching

4   event in that usage session for each usage event of the sequence differs from the

5   occurrence time of the matching event in that usage session for the next usage event of

6   the sequence by less than a determined amount.

1   63.   The method of claim 49 wherein the usage events in the usage

2   sessions reflect interactions by users with an executing application.

1   64.   The method of claim 49 wherein the usage events in the usage

2   sessions reflect interactions by users with a provided service.

1   65.   The method of claim 49 wherein the usage events in the usage

2   sessions reflect acquisitions of provided items by remote users.

1   66.   The method of claim 49 wherein the usage events in the usage

2   sessions reflect Web pages provided to users.

1   67.   The method of claim 49 wherein at least some of the usage events in

2   the usage sessions are each related to an item having an associated attribute, and wherein

3   during the determining of the sequences of usage events only sequence events in the

4   usage sessions having a related item whose associated attribute is a specified attribute are

5   allowed to be matching events.

68.   The method of claim 67 wherein each item has an associated attribute that specifies a category indicating a type of the item, and wherein the specified attribute is a particular category.

69.   The method of claim 49 wherein at least some of the usage events in the usage sessions are of a specified type, and wherein during the determining of the sequences of usage events only sequence events in the usage sessions of the specified type are allowed to be matching events.

70.   The method of claim 49 wherein each usage event in the usage sessions is related to an item having an associated attribute, and wherein during the determining of the sequences of usage events each sequence event in each usage session is represented by the associated attribute of the related item for the sequence event in such a manner that each determined sequence of usage events represents a sequence of the attributes that are associated with the items related to the usage events in the sequence and that is determined to frequently be present in the sequences of attributes representing the sequences of events in the usage sessions.

71.   The method of claim 70 wherein each of the attributes specifies a category indicating a type of the associated item.

72.   The method of claim 49 wherein each of the usage events in the usage sessions are of a specified class, and wherein during the determining of the sequences of usage events each sequence event in each usage session is represented by the specified class of the sequence event in such a manner that each determined sequence of usage events represents a sequence of the classes of the usage events in the sequence, the represented sequence of classes determined to frequently be present in the sequences of classes representing the sequences of events in the usage sessions.

1   73.   The method of claim 49 wherein each usage session reflects one or

2 more related interaction events by a user.

1   74.   The method of claim 49 wherein each usage session reflects one or

2 more related Web pages provided to a user during a browsing session.

1   75.   The method of claim 49 wherein each usage session reflects one or

2 more related Web pages provided to a user during multiple browsing sessions.

1   76.   The method of claim 49 wherein each usage session reflects one or

2 more items specified by a remote user.

1   77.   The method of claim 49 wherein each usage session reflects usage

2 events by multiple users during a specified period of time.

1   78.   The method of claim 49 wherein when a sequence of usage events is

2 present multiple times in a usage session, only one of the presences is counted when

3 determining if the sequence is present more than the threshold.

1   79.   A computer-readable medium whose contents cause a computing

2 device to identify ordered sequences of usage events that frequently occur during usage

3 sessions by:

4   for each of multiple usage sessions, receiving an indication of an ordered

5 sequence of usage events that occurred during the usage session;

6   for each of multiple sequences of usage events, determining that the

7 sequence of usage events is present in the indicated usage sessions more times than a

8 threshold, a sequence of usage events determined to be present in a usage session when

9 each usage event of the sequence has a distinct matching event in the ordered sequence

10 for the usage session such that the matching event for each usage event in the sequence

11   appears in the ordered sequence before the matching event for a next usage event in the

12   sequence; and

13   providing indications of the determined sequences of usage events.

1   80.   The computer-readable medium of claim 79 wherein the contents

2   further cause the computer device to receive at least one log from a customer reflecting

3   usage by multiple users of a service provided by that customer, wherein the indications of

4   the ordered sequences of related usage events for the multiple usage sessions are received

5   by identifying the multiple usage sessions from the received log, and wherein the

6   indications of the determined sequences of usage events are provided to a representative

7   of the customer.

1   81.   The computer-readable medium of claim 80 wherein the provided

2   indications are part of a generated report.

1   82.   The computer-readable medium of claim 79 including receiving an

2   instruction from a remote user, and presenting the determined sequences of usage events

3   to the remote user in response.

1   83.   The computer-readable medium of claim 79 wherein the multiple

2   sequences of usage events are each present in usage sessions in a non-consecutive order

3   in at least some of the more times, a sequence of usage events determined to be present in

4   a usage session in a non-consecutive order when two of the matching events for usage

5   events of the sequence are separated in the ordered sequence of the usage session by at

6   least one intervening usage event that is not part of the sequence of usage events.

1   84.   The computer-readable medium of claim 79 wherein when a

2   sequence of usage events is present multiple times in a usage session, only one of the

3   presences is counted when determining if the sequence is present more than the threshold.

1   85.   The computer-readable medium of claim 79 wherein the computer-

2   readable medium is a data transmission medium transmitting a generated data signal

3   containing the contents.

1   86.   The computer-readable medium of claim 79 wherein the computer-

2   readable medium is a memory of a computer system.

1   87.   A computing device for identifying ordered sequences of usage

2   events that frequently occur during usage sessions, comprising:

3   an interaction event supplier component capable of receiving for each of

4   multiple usage sessions an indication of an ordered sequence of usage events that

5   occurred during the usage session; and

6   a frequent sequence analyzer component capable of, for each of multiple

7   sequences of usage events, determining that the sequence of usage events is present in the

8   indicated usage sessions more times than a threshold, a sequence of usage events

9   determined to be present in a usage session when each usage event of the sequence has a

10 distinct matching event in the ordered sequence for the usage session such that the

11 matching event for each usage event in the sequence appears in the ordered sequence

12 before the matching event for a next usage event in the sequence.

1   88.   The computing device of claim 87 including a frequent sequence

2   indication generator component capable of providing indications of the determined

3   sequences of usage events.

1   89.   The computing device of claim 87 wherein the interaction event

2   supplier component further receives at least one log from a customer reflecting usage by

3   multiple users of a service provided by that customer, wherein the indications of the

4   ordered sequences of related usage events for the multiple usage sessions are received by

5   identifying the multiple usage sessions from the received log, and wherein the indications

6   of the determined sequences of usage events are provided to a representative of the

7   customer.

1   90.   The computing device of claim 87 further comprising a component

2   that is capable of receiving an instruction from a remote user and of presenting the

3   determined sequences of usage events to the remote user in response.

1   91.   The computing device of claim 87 wherein the interaction event

2   supplier component and the frequent sequence analyzer component are executing in

3   memory of the computing device.

1   92.   A computer system for identifying ordered sequences of usage

2   events that frequently occur during usage sessions, comprising:

3   means for, for each of multiple usage sessions, receiving an indication of an

4   ordered sequence of usage events that occurred during the usage session;

5   means for, for each of multiple sequences of usage events, determining that

6   the sequence of usage events is present in the indicated usage sessions more times than a

7   threshold, a sequence of usage events determined to be present in a usage session when

8   each usage event of the sequence has a distinct matching event in the ordered sequence

9   for the usage session such that the matching event for each usage event in the sequence

10   appears in the ordered sequence before the matching event for a next usage event in the

11   sequence; and

12   means for providing indications of the determined sequences of usage

13   events.

93.   A computer-readable generated data signal transmitted via a transmission medium, the generated data signal having encoded contents that cause a computer system to identify ordered sequences of usage events that frequently occur during usage sessions by:

for each of multiple usage sessions, receiving an indication of an ordered sequence of usage events that occurred during the usage session;

for each of multiple sequences of usage events, determining that the sequence of usage events is present in the indicated usage sessions more times than a threshold, a sequence of usage events determined to be present in a usage session when each usage event of the sequence has a distinct matching event in the ordered sequence for the usage session such that the matching event for each usage event in the sequence appears in the ordered sequence before the matching event for a next usage event in the sequence; and

providing indications of the determined sequences of usage events.

94.   The computer-readable generated data signal of claim 93 wherein the generated data signal is encoded as part of a carrier wave.

95.   A data structure stored in a memory for use in identifying frequent sequences of interaction events, the data structure containing a root element and a plurality of other elements, a first level of the data structure having elements that are children elements of the root element, each of multiple other levels having elements that are children elements of elements of a previous level, each of the plurality of other elements representing one of the interaction events and being associated with a candidate sequence,

so that after the elements are scored while validating the data structure against groups of related interaction events, the candidate sequences that are associated with elements whose score is above a threshold can be selected as the identified frequent sequences of interaction events.

1   96.   The data structure of claim 95 wherein each element other than the

2 root element has an associated sequential path of elements between the root element and

3 that element, a first element in each sequential path being a child element of the root

4 element, each element in each sequential path other than the first element being a child

5 element of the previous element in the sequential path, and wherein the candidate

6 sequence of interaction events that is associated with each node other than the root node

7 consists of a sequence of the interaction events represented by the elements in the path

8 associated with that node followed by the interaction event represented by that node.

1   97.   The data structure of claim 95 including, for each interaction event

2 represented by a child element of the root element, a linked list data structure that is

3 associated with that interaction event such that the linked list data structure includes

4 entries for at least one of the other elements,

5 so that as the data structure is validated against one of the groups of interaction events by

6 selecting in order each interaction event in the group, if the selected interaction event is

7 one of the interaction events having an associated linked list data structure, the score of

8 each of the elements having entries in the associated linked list data structure can be

9 incremented.

1   98.   The data structure of claim 95 wherein at least some of the elements

2 have associated tokens,

3 so that as the data structure is validated against one of the groups of interaction events by

4 selecting in order each interaction event in the group, the score can be incremented for

5 each element that represents the selected interaction event, that lacks an associated token,

6 and that is a child element of an element with an associated token.

7